

# **TAC KBP Sentiment Slot Filling Guidelines**

Version 1.0, July 29, 2013

**Linguistic Data Consortium**

## Table of Contents

|   |          |
|---|----------|
| <b>1. Introduction .....</b>                      | <b>1</b> |
| <b>2. Sentiment Slots and Fillers .....</b>       | <b>1</b> |
| 2.1 pos-towards .....                             | 1        |
| 2.2 pos-from.....                                 | 2        |
| 2.3 neg-towards .....                             | 2        |
| 2.4 neg-from.....                                 | 2        |
| <b>3. Slot Content.....</b>                       | <b>3</b> |
| <b>4. Slot Quantity.....</b>                      | <b>3</b> |
| <b>5. Unique Fillers .....</b>                    | <b>3</b> |
| <b>6. Annotating Sentiment Query Answers.....</b> | <b>4</b> |
| 6.1 Basis for sentiment .....                     | 4        |
| 6.2 Post authors and bloggers .....               | 4        |
| 6.3 Different sentiment expressions.....          | 4        |
| 6.4 Out of scope sentiment expressions .....      | 5        |

## 1. Introduction

Text Analysis Conference (TAC) is a series of workshops organized by the National Institute of Standards and Technology (NIST). TAC was developed to encourage research in natural language processing (NLP) and related applications by providing a large test collection, common evaluation procedures, and a forum for researchers to share their results. The Knowledge Base Population (KBP) track of TAC aims to develop systems that can determine whether or not entities have an existing Wikipedia page, extract information about those entities from web and newswire texts, and use the extracted information to populate an existing knowledge base.

The goal of Sentiment Slot Filling is to use corpora to collect information regarding sentiment expressed towards or by an entity. Sentiment is defined here as a positive or negative emotion, evaluation, or judgment.

To support this task, we will develop sentiment queries and annotation of answers to those queries. Queries include a query entity and a sentiment slot that indicates both query polarity (whether the sentiment is positive or negative) and directionality (whether the query entity is the holder or the target of the sentiment), and fillers will be the entities (PER, ORG, or GPE) at whom the query entity is directing the sentiment, or who hold sentiment toward the query entity (depending on the query).

## 2. Sentiment Slots and Fillers

There are just four slots in the TAC KBP sentiment slot filling task which will be used in query development and annotation:

### 2.1 pos-towards

Query entity holds positive sentiment towards filler entity (“likes”, “is hopeful about”, etc.). In the triple <sentiment holder, sentiment, sentiment target>, the fillers are the sentiment targets.

| Slot               | Query Entity    | Document Text  | Filler Entity |
|--------------------|-----------------|--|---------------|
| <i>pos-towards</i> | Whoopi Goldberg | “I admired Maria’s defense of Arnold even if he doesn’t deserve it and will have hell to pay when they get home alone,” said Goldberg. | Maria Shriver |
| <i>pos-towards</i> | Steven Flaherty | Flaherty also said NBC’s cooperation with authorities was appreciated by all involved.   | NBC           |

## 2.2 pos-from

Query entity is target of positive sentiment from filler entity (“is liked by”, “was hoped for by”, etc.). In the triple <sentiment holder, sentiment, sentiment target>, the fillers are the sentiment holders.

| Slot            | Query Entity   | Document Text  | Filler Entity                      |
|-----------------|----------------|--|------------------------------------|
| <i>pos-from</i> | NBC            | Flaherty also said NBC’s cooperation with authorities was appreciated by all involved. | Steven Flaherty                    |
| <i>pos-from</i> | Oscar Peterson | <u>cicerone imposter</u><br>The sounds Peterson produces on the keys are beyond words. | cicerone impostor<br>[post author] |

## 2.3 neg-towards

Query entity holds negative sentiment towards filler entity (“dislikes”, “is skeptical about”, etc.). In the triple <sentiment holder, sentiment, sentiment target>, the fillers are the sentiment targets.

| Slot               | Query Entity            | Document Text  | Filler Entity |
|--------------------|-------------------------|--|---------------|
| <i>neg-towards</i> | National Beverage Corp. | Still, the city of Lansing is the capital of Michigan, so my company [National Beverage Corp.] expected more.                            | Lansing       |
| <i>neg-towards</i> | Steve Jobs              | Steve Jobs even has the audacity to criticize Gates at every opportunity, even though Bill Gates is intellectually far superior to Jobs. | Bill Gates    |

## 2.4 neg-from

Query entity is target of negative sentiment from filler entity (“is disliked by”, etc.). In the triple <sentiment holder, sentiment, sentiment target>, the fillers are the sentiment holders.

| Slot            | Query Entity | Document Text  | Filler Entity |
|-----------------|--------------|--|---------------|
| <i>neg-from</i> | Bill Gates   | Steve Jobs even has the audacity to criticize Gates at every opportunity, even though Bill Gates is intellectually far superior to Jobs. | Steve Jobs    |

|                 |                             |  |                                   |
|-----------------|-----------------------------|--|-----------------------------------|
| <i>neg-from</i> | Central Intelligence Agency | <u>whitehall</u><br>The CIA has been around since WW2 and it's big on Central and short on intelligence. | whitehall<br><i>[post author]</i> |
|-----------------|-----------------------------|--|-----------------------------------|

### 3. Slot Content

All four of the slots for Sentiment Slot Filling are name slots, meaning that they are required to be filled by the name of a person, organization, or geo-political entity (GPE):

- **Person Entities (PER)** – PERs are limited to individual humans. Groups of people (including families) are not valid person entities.
- **Organization Entities (ORG)** – ORGs are corporations, agencies, and other groups of people defined by an established organizational structure. Note that musical groups are considered to be organizations but individual artists (e.g. Britney Spears) are considered persons. Programs or projects should not be considered organizations and different iterations of the same organization (e.g., the 111th U.S. Congress and the 112th U.S. Congress) should not be considered as distinct entities.
- **Geo-political Entities (GPE)** – Generally speaking, GPEs are composite entities comprised of a government, a physical location, and a population, with common types including countries, states, provinces, counties, cities, and towns. Note, however, that for the purposes of TAC KBP, all top-level governments of GPEs should also be categorized as GPEs, not as ORGs.

Regions like “the southeast US” are not GPEs because, while they have the physical location and population qualities, they do not have their own government. Given the text “southeast Texas”, only “Texas” could be annotated as GPE, as southeast Texas has neither its own government nor a defined location.

### 4. Slot Quantity

All four of the slots for Sentiment Slot Filling are list value, meaning that they can take multiple fillers as they are likely to have more than one correct answer in the source data.

### 5. Unique Fillers

Multiple fillers returned for the same query should refer to distinct individuals. It is not sufficient that slot filler entity strings be distinct; they must refer to distinct individuals.

For example, if the query includes {Hillary Clinton, pos-towards} (the sentiment holder is Hillary Clinton with positive sentiment towards the filler), and we find

both “William Clinton” and “Bill Clinton” as potential fillers, we should return just one of these.

**NOTE:** Similarly, the same entity should not be repeated as a filler for a query, even if the same polarity of sentiment is held toward or about an entity for different reasons.

For example, although it is possible that Hillary Clinton may feel “pos-towards” William Jefferson Clinton on many separate occasions, we should only annotate one of these instances of her holding positive sentiment toward him.

## **6. Annotating Sentiment Query Answers**

### **6.1 Basis for sentiment**

The most important thing to note is that sentiment may be directed toward an entity based on direct evaluation of an entity (e.g. “Kentucky doesn't like Mitch McConnell”) or may be directed to an entity based on actions that the entity took (e.g. “Kentucky doesn't like Mitch McConnell’s stance on gun control”).

### **6.2 Post authors and bloggers**

In the discussion forum and weblog data, post authors and bloggers may be used as query entities, or returned as filler entities, though they should only be used as query entities if they can be positively identified and thus either linked to the KB or marked as NIL.

### **6.3 Different sentiment expressions**

Simple/newswire type sentiment expression (use this less)

Henderson said Smith was "very smart" and a "great businessman."

Complex/discussion forum type sentiment expression (use this more)

#### **Debra Law**

Mortkat, the less-than-masterful king of feeble arguments and feeble attempts at deflection, accuses others of his own flaws, AGAIN.

#### **pixee**

1.Perth (south west coast) has a lower crime rate and quieter lifestyle than some of the east coast cities, it's also said to have the best beaches (the gold coast in Queensland is nice too).

**TIP:** Search in the discussion forum and blog data – this will have more expressions of sentiment than newswire.

## 6.4 Out of scope sentiment expressions

Complex uses of sarcasm (these are too hard for systems)

### Liability

Call in Halliburton! Send down some fucking pipes to release the super-heated pressure and USE it to generate steam, electricity, heating, etc. But mostly to release the pressure. Oh, and to piss off the Arab Oil cartel. I've now officially heard ENOUGH of this "SAVE the WHALES" malarkey. Fuck the whales. Let's save North America from the caldera under Yellowstone. I mean, come ON! Are we gonna let some little super volcano defeat us? Pfft